Parent case Child case 10/068,080 Seq Id #1

****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

```
Run on:          Wed Mar  8 10:23:17 2000;  MasPar time 10.58 Seconds
                                            601.178 Million cell updates/sec
Tabular output not generated.

Title:           >US-09-163-445-1
Description:     (1-213) from US09163445.pep
Perfect Score:   1544
Sequence:        1 MITIVLLILAYLLGSIPSGL..........NKTENLVPWGLNLTHQDPKK 213

Scoring table:   PAM 150
                 Gap 11

Searched:        82229 seqs, 29864866 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        swiss-prot38
                   1:swissprot

Statistics:      Mean 48.353;  Variance 138.901;  scale 0.348
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	1536	99.5	213	1	YPAE_STRPN	HYPOTHETICAL 23.0 KD P	6.13e-198
2	616	39.9	193	1	YNES_BACSU	HYPOTHETICAL 21.0 KD P	4.10e-65
3	460	29.8	192	1	Y676_AQUAE	HYPOTHETICAL PROTEIN A	1.59e-43
4	354	22.9	222	1	YJ73_SYNY3	HYPOTHETICAL 23.5 KD P	2.75e-29
5	297	19.2	205	1	YGIH_ECOLI	HYPOTHETICAL 22.2 KD P	7.28e-22
6	294	19.0	199	1	YGIH_HAEIN	HYPOTHETICAL PROTEIN H	1.76e-21
7	256	16.6	262	1	YF09_HELPY	HYPOTHETICAL PROTEIN H	1.13e-16
8	188	12.2	239	1	Y247_MYCGE	HYPOTHETICAL PROTEIN M	1.75e-08
9	158	10.2	239	1	Y247_MYCPN	HYPOTHETICAL PROTEIN M	4.06e-05
10	145	9.4	820	1	NAH1_MOUSE	SODIUM/HYDROGEN EXCHAN	9.97e-04
11	142	9.2	817	1	NAH1_BOVIN	SODIUM/HYDROGEN EXCHAN	2.05e-03
12	142	9.2	822	1	NAH1_CRIGR	SODIUM/HYDROGEN EXCHAN	2.05e-03
13	141	9.1	816	1	NAH1_RABIT	SODIUM/HYDROGEN EXCHAN	2.61e-03
14	140	9.1	818	1	NAH1_PIG	SODIUM/HYDROGEN EXCHAN	3.31e-03
15	139	9.0	820	1	NAH1_RAT	SODIUM/HYDROGEN EXCHAN	4.20e-03
16	137	8.9	827	1	SP2E_BACSU	STAGE II SPORULATION P	6.74e-03
17	133	8.6	429	1	URAA_ECOLI	URACIL PERMEASE (URACI	1.72e-02
18	132	8.5	540	1	YM8M_YEAST	HYPOTHETICAL 59.6 KD P	2.17e-02
19	129	8.4	402	1	YHJX_ECOLI	HYPOTHETICAL 43.0 KD P	4.33e-02
20	130	8.4	435	1	Y609_METJA	HYPOTHETICAL PROTEIN M	3.44e-02
21	127	8.2	1049	1	ACRB_ECOLI	ACRIFLAVIN RESISTANCE	6.84e-02
22	125	8.1	669	1	YH06_HAEIN	PROBABLE TRANSPORT PRO	1.08e-01
23	125	8.1	815	1	NAH1_HUMAN	SODIUM/HYDROGEN EXCHAN	1.08e-01
24	122	7.9	623	1	PTR2_CANAL	PEPTIDE TRANSPORTER PT	2.11e-01
25	122	7.9	648	1	CCMF_HAEIN	CYTOCHROME C-TYPE BIOG	2.11e-01
26	120	7.8	173	1	NU6M_SALSA	NADH-UBIQUINONE OXIDOR	3.29e-01
27	121	7.8	440	1	ATOE_ECOLI	SHORT-CHAIN FATTY ACID	2.63e-01
28	119	7.7	118	1	Y759_METJA	HYPOTHETICAL PROTEIN M	4.10e-01
29	119	7.7	277	1	Y233_METJA	HYPOTHETICAL PROTEIN M	4.10e-01
30	119	7.7	446	1	CITN_SALDU	CITRATE-SODIUM SYMPORT	4.10e-01
31	119	7.7	446	1	CITN_SALPU	CITRATE-SODIUM SYMPORT	4.10e-01
32	119	7.7	741	1	YF01_MYCPN	HYPOTHETICAL 85.0 KD P	4.10e-01
33	118	7.6	113	1	Y406_MYCGE	HYPOTHETICAL PROTEIN M	5.10e-01
34	118	7.6	323	1	YJFF_ECOLI	HYPOTHETICAL ABC TRANS	5.10e-01
35	118	7.6	328	1	GPT_SULAC	PUTATIVE UDP-N-ACETYLG	5.10e-01
36	118	7.6	585	1	SP2E_BACME	STAGE II SPORULATION P	5.10e-01
37	115	7.4	311	1	YXXF_BACSU	HYPOTHETICAL 34.3 KD P	9.78e-01
38	115	7.4	326	1	MRAY_STRPN	PHOSPHO-N-ACETYLMURAMO	9.78e-01
39	115	7.4	400	1	TYRP_HAEIN	TYROSINE-SPECIFIC TRAN	9.78e-01
40	115	7.4	424	1	FLO3_CAEEL	FOLATE-LIKE TRANSPORTE	9.78e-01
41	114	7.4	440	1	CAPE_STAAU	CAPE PROTEIN.	1.21e+00
42	115	7.4	675	1	NU5M_ACACA	NADH-UBIQUINONE OXIDOR	9.78e-01
43	114	7.4	787	1	SP3E_BACSU	STAGE III SPORULATION	1.21e+00
44	115	7.4	877	1	PMT1_CANAL	DOLICHYL-PHOSPHATE-MAN	9.78e-01
45	114	7.4	1037	1	YHIV_ECOLI	HYPOTHETICAL 111.5 KD	1.21e+00

ALIGNMENTS

```
RESULT    1
ID   YPAE_STRPN     STANDARD;      PRT;   213 AA.
AC   Q54916;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   HYPOTHETICAL 23.0 KD PROTEIN IN PARE 5'REGION (ORF2).
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=7785;
RX   MEDLINE; 96312346.
RA   PAN X., FISHER M.;
RT   "Cloning and characterization of the parC and parE genes of
RT   Streptococcus pneumoniae encoding DNA topoisomerase IV: role in
RT   fluoroquinolone resistance.";
RL   J. Bacteriol. 178:4060-4069(1996).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z67739; CAA91549.1; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM      2     22       POTENTIAL.
FT   TRANSMEM     70     90       POTENTIAL.
FT   TRANSMEM    112    132       POTENTIAL.
FT   TRANSMEM    143    163       POTENTIAL.
FT   TRANSMEM    164    184       POTENTIAL.
SQ   SEQUENCE   213 AA;  22988 MW;  4C2842DB CRC32;

  Query Match            99.5%;  Score 1536;  DB 1;  Length 213;
  Best Local Similarity  99.5%;  Pred. No. 6.13e-198;
  Matches  212;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Db        1 MITIVLLILAYLLGSIPSGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVID 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        1 MITIVLLILAYLLGSIPSGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVID 60
```

```
Db     61 FFKGTLATLLPIIFHLQGVSPLIFGLLAVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIF 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 FFKGTLATLLPIIFHLQGVSPLIFGLLAVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIF 120

Db    121 CLYLAIIFFGALYLGSMISLSSVTASIAAVIGVLLFPLFGFILSNYDFLFIAIILALASL 180
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Qy    121 CLYLAIIFFGALYLGSMISLSSVTASIAAVIGVLLFPLFGFILSNYDSLFIAIILALASL 180

Db    181 IIIRHKDNIARIKNKTENLVPWGLNLTHQDPKK 213
          ||||||||||||||||||||||||||||||||||
Qy    181 IIIRHKDNIARIKNKTENLVPWGLNLTHQDPKK 213


RESULT    2
ID    YNES_BACSU     STANDARD;      PRT;   193 AA.
AC    Q45064;
DT    15-JUL-1999 (Rel. 38, Created)
DT    15-JUL-1999 (Rel. 38, Last sequence update)
DT    15-JUL-1999 (Rel. 38, Last annotation update)
DE    HYPOTHETICAL 21.0 KD PROTEIN IN TLP-GRLB INTERGENIC REGION.
GN    YNES.
OS    Bacillus subtilis.
OC    Bacteria; Firmicutes; Bacillus/Clostridium group;
OC    Bacillus/Staphylococcus group; Bacillus.
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=168;
RA    ROSE M., ENTIAN K.;
RL    Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC    -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC    -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; Z73234; CAA97604.1; -.
DR    EMBL; Z99113; CAB13690.1; -.
DR    SUBTILIST; BG11826; YNES.
KW    Hypothetical protein; Transmembrane.
FT    TRANSMEM      2     22       POTENTIAL.
FT    TRANSMEM     54     74       POTENTIAL.
FT    TRANSMEM     76     96       POTENTIAL.
FT    TRANSMEM    111    131       POTENTIAL.
FT    TRANSMEM    153    173       POTENTIAL.
SQ    SEQUENCE   193 AA;  20966 MW;  BC042392 CRC32;

  Query Match           39.9%;  Score 616;  DB 1;  Length 193;
  Best Local Similarity  50.0%;  Pred. No. 4.10e-65;
  Matches   101;  Conservative   45;  Mismatches  46;  Indels  10;  Gaps   6;

Db      1 MLIALLIILAYLIGSIPSGLIVGKLAKGIDIREHGSGNLGATNAFRTLGVKAGSVVIAGD 60
          |:  :|:|||||:|||||||  :| :   |::|||||||  |:||:|| || |||   :  |
Qy      1 MITIVLLILAYLLGSIPSGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVID 60

Db     61 ILKGTLATALPFLMHVD-IHPLLAGVFAVLGHVFPIFAKFKGGKAVATSGGVLLFYAPLL 119
          ::||||||  ||:: |:: : ||: |::||:|| ||||| ||||||||||:||:: :||::
Qy     61 FFKGTLATLLPIIFHLQGVSPLIFGLLAVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIF 120

Db    120 FITMVAVFFIFLYLTKFVSLSSMLTGIYTVI----YS-F-FVHDTY--LLIVVTL-LTIF 170
           : :  :||  |||   :|||| ::| :||    :: | |:   |  |:| : | |: :
Qy    121 CLYLAIIFFGALYLGSMISLSSVTASIAAVIGVLLFPLFGFILSNYDSLFIAIILALASL 180

Db    171 VIYRHRANIKRIINKTEPKVKW 192
          :| ||: || || ||||  | |
Qy    181 IIIRHKDNIARIKNKTENLVPW 202


RESULT    3
ID    Y676_AQUAE     STANDARD;      PRT;   192 AA.
AC    O66905;
DT    15-JUL-1999 (Rel. 38, Created)
DT    15-JUL-1999 (Rel. 38, Last sequence update)
DT    15-JUL-1999 (Rel. 38, Last annotation update)
DE    HYPOTHETICAL PROTEIN AQ_676.
GN    AQ_676.
OS    Aquifex aeolicus.
OC    Bacteria; Aquificales; Aquificaceae; Aquifex.
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=VF5;
RX    MEDLINE; 98196666.
RA    DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA    GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA    FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT    "The complete genome of the hyperthermophilic bacterium Aquifex
RT    aeolicus.";
RL    Nature 392:353-358(1998).
CC    -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC    -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; AE000701; AAC06869.1; -.
KW    Hypothetical protein; Transmembrane.
FT    TRANSMEM      3     23       POTENTIAL.
FT    TRANSMEM     51     71       POTENTIAL.
FT    TRANSMEM     80    100       POTENTIAL.
FT    TRANSMEM    112    132       POTENTIAL.
FT    TRANSMEM    149    169       POTENTIAL.
SQ    SEQUENCE   192 AA;  20940 MW;  CB7A57E2 CRC32;

  Query Match           29.8%;  Score 460;  DB 1;  Length 192;
  Best Local Similarity  42.6%;  Pred. No. 1.59e-43;
  Matches    83;  Conservative   55;  Mismatches  43;  Indels  14;  Gaps  13;

Db      1 MKALFLVIFAYLLGSITFGEVIAKLK-GVDLRNVGSGNVGATNVTRALGKKYGVLVFFLD 59
          | :: |:|:|||||||  |  |: :   ::||: |||| |:||  | |||| |:  | :|
Qy      1 MITIVLLILAYLLGSIPSGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVID 60

Db     60 FLKGFIPALIAVKSFGIDS-WVLTFTGLASVLGHMYPVFFGFKGGKGVATALGVVFAVSP 118
          |:|| :::|:::  | :::   | | || :|:|| :|:| |||||:|||: ||:|: :|
Qy     61 FFKGTLATLLPII-FHLQGVSPLIF-GLLAVIGHTFPIFAGFKGGKAVATSAGVIFGFAP 118

Db    119 SVALFSFLVWLG-IFLWKRYVSLASITATISA-F-LFLF-VAGYPV-N--VLFMAIVIG- 170
            |:  :: :| ::|    :||:|:||:|:| : ::|| : |: : |   ||:||:::
Qy    119 IFCLYLAIIFFGALYL-GSMISLSSVTASIAAVIGVLLFPLFGFILSNYDSLFIAIILAL 177

Db    171 A-LIIYRHRENINRL 184
          | ||| ||::|| |:
Qy    178 ASLIIIRHKDNIARI 192


RESULT    4
ID    YJ73_SYNY3     STANDARD;      PRT;   222 AA.
AC    P73933;
DT    15-JUL-1999 (Rel. 38, Created)
DT    15-JUL-1999 (Rel. 38, Last sequence update)
DT    15-JUL-1999 (Rel. 38, Last annotation update)
DE    HYPOTHETICAL 23.5 KD PROTEIN SLL1973.
GN    SLL1973.
OS    Synechocystis sp. (strain PCC 6803).
OC    Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE; 97061201.
```

```
RA   KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA   MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA   HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA   SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA   TABATA S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; D90910; BAA17999.1; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM      4     24       POTENTIAL.
FT   TRANSMEM     56     76       POTENTIAL.
FT   TRANSMEM     87    107       POTENTIAL.
FT   TRANSMEM    130    150       POTENTIAL.
FT   TRANSMEM    153    173       POTENTIAL.
SQ   SEQUENCE   222 AA;  23534 MW;  D5B6F852 CRC32;

  Query Match           22.9%;  Score 354;  DB 1;  Length 222;
  Best Local Similarity  39.2%;  Pred. No. 2.75e-29;
  Matches    82;  Conservative   56;  Mismatches  50;  Indels  21;  Gaps  17;

Db       6 LLCLCLLLLITYLMGSIPTGYLAGKLLLGIDIREHGSKSTGATNVFRTLGKPAAIAVLAID 65
           :: : ||:::||:||||:|   | ::: |::|||| :||:|| || ||| |::| : ||
Qy       1 MITIVLLILAYLLGSIPSGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVID 60

Db      66 ISKGVMAVALVRAIYSGDWLPALPAAWQNWLTLGVAIAVVLGHSKSIFLKFSGGKSVATS 125
           : || :| :|:  |:    | :: ::    | :|: :||: ||: :||  | |||:||||
Qy      61 FFKGTLA-TLLPIIF--H-LQGV-SP----LIFGL-LAVI-GHTFPIFAGFKGGKAVATS 109

Db     126 LGVLFMLNIWLALGTLAT-FLTVIFFTRIVSLSSIVAAIA-VNGIAL-AL-Q-LPPPYLA 180
            ||:| :   : |  ||  |:  :::  ::|||| |:|| | |: | :|   : : | :
Qy     110 AGVIFGFAPIFCL-YLAIIFFGALYLGSMISLSSVTASIAAVIGVLLFPLFGFILSNYDS 168

Db     181 -F-TF-LA-GMYVIVRHRTNIERILQGTE 205
            | :: || :  :|:||: || ||   ||
Qy     169 LFIAIILALASLIIIRHKDNIARIKNKTE 197


RESULT    5
ID   YGIH_ECOLI     STANDARD;      PRT;   205 AA.
AC   P31056;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   HYPOTHETICAL 22.2 KD PROTEIN IN BACA-TTDA INTERGENIC REGION (O205).
GN   YGIH.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE; 97426617.
RA   BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA   RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA   GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA   MAU B., SHAO Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE OF 1-204 FROM N.A.
RC   STRAIN=JM83;
RX   MEDLINE; 93285992.
RA   CAIN B.D., NORTON P.J., EUBANKS W., NICK H.S., ALLEN C.M.;
RT   "Amplification of the bacA gene confers bacitracin resistance to
RT   Escherichia coli.";
RL   J. Bacteriol. 175:3784-3789(1993).
RN   [3]
RP   IDENTIFICATION.
RX   MEDLINE; 95075659.
RA   BORODOVSKY M., RUDD K.E., KOONIN E.V.;
RT   "Intrinsic and extrinsic approaches for detecting genes in a
RT   bacterial genome.";
RL   Nucleic Acids Res. 22:4756-4767(1994).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY. STRONG, TO H.INFLUENZAE
CC       HI0266.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U28379; AAA89139.1; -.
DR   EMBL; AE000387; AAC76095.1; -.
DR   EMBL; L12966; -; NOT_ANNOTATED_CDS.
DR   ECOGENE; EG11674; YGIH.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM      4     24       POTENTIAL.
FT   TRANSMEM     53     73       POTENTIAL.
FT   TRANSMEM     82    102       POTENTIAL.
FT   TRANSMEM    112    132       POTENTIAL.
FT   TRANSMEM    138    158       POTENTIAL.
SQ   SEQUENCE   205 AA;  22193 MW;  FE3A2166 CRC32;

  Query Match           19.2%;  Score 297;  DB 1;  Length 205;
  Best Local Similarity  36.3%;  Pred. No. 7.28e-22;
  Matches    74;  Conservative   56;  Mismatches  57;  Indels  17;  Gaps  13;

Db       4 IAPGMILIAYLCGSISSAILVCRLCGLPDPRTSGSGNPGATNVLRIGGKGAAVAVLIFDV 63
           |:  ::::||| |||:|:: : ::     : |  |||| |:|| :|| || |::| :::|
Qy       2 ITIVLLILAYLLGSIPSGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVIDF 61

Db      64 LFGMLP-VWGA-YEL-GVSPFWLGLIAIAACLGHIWPVFFGFKGGKGVATAFGAI---AP 117
           :|| |: :    : | ||||: :||:|:   :|| · |:| |||||:|||: | |   ||
Qy      62 FFGTLATLLPIIFHLQGVSPLIFGLLAV---IGHTFPIFAGFKGGKAVATSAGVIFGFAP 118

Db     118 IG-WDLTGVMAGT-WLLTVL-LSGYS-SLGAIVSALIAP-F-YVWW-FKPQFTFPVSMLS 170
           |    |: :  |:  | ::: ||: : |::|::: |: | | ::   : : |   :  |:
Qy     119 IFCLYLAIIFFGALYLGSMISLSSVTASIAAVIGVLLFPLFGFILSNYDSLFIAIILALA 178

Db     171 CLILLRHHDNIQRLWRRQETKI-W 193
            ||::|| ||| |:  : |  : |
Qy     179 SLIIIRHKDNIARIKNKTENLVPW 202


RESULT    6
ID   YGIH_HAEIN     STANDARD;      PRT;   199 AA.
AC   P44603;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   HYPOTHETICAL PROTEIN HI0266.
GN   HI0266.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20;
RX   MEDLINE; 95350630.
```

```
RA   FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA   KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA   MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA   SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
RA   WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA   UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA   FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA   GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA   VENTER J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY. STRONG, TO E.COLI YGIH.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32713; AAC21932.1; -.
DR   TIGR; HI0266; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM      4     24       POTENTIAL.
FT   TRANSMEM     56     76       POTENTIAL.
FT   TRANSMEM     80    100       POTENTIAL.
FT   TRANSMEM    115    135       POTENTIAL.
FT   TRANSMEM    138    158       POTENTIAL.
SQ   SEQUENCE   199 AA;  22113 MW;  31409562 CRC32;

  Query Match           19.0%;  Score 294;  DB 1;  Length 199;
  Best Local Similarity  35.3%;  Pred. No. 1.76e-21;
  Matches    72;  Conservative   52;  Mismatches  67;  Indels  13;  Gaps  10;

Db       3 LFALFYMLFAYLLGSISSAILICRIAGLPDPRQNGSHNPGATNVLRIGNRKSALAVLIFD 62
           ::::  :::|||||||:|:: | ::    : |::|| | |:|| :||  :|:::| :::|
Qy       1 MITIVLLILAYLLGSIPSGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVID 60

Db      63 MLKGMIPVWAGYYLGLTQFELGMVALGACLGHIFPIFFQFKGGKGVATAFGAI---API- 118
             :|| ::      : |      : :| | :|| |||| |||||:|||: | |   |||
Qy      61 FFKGTLATLLPIIFHLQGVSPLIFGLLAVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIF 120

Db     119 SWAVAGSMFGT-WI--FVFLVSGYSSLSAVISALLVP-F-YVWW-FKPEF-TFPVALVCC 171
           :|   ||:  :   : | |  :|::|||: || | | ::   : : | :: :||
Qy     121 CLYLAIIFFGALYLGSMISLSSVTASIAAVIGVLLFPLFGFILSNYDSLFIAIILALAS- 179

Db     172 LLIYRHHDNIQRLWRGQEDKV-WA 194
           |:| || ||| |:    |: | |:
Qy     180 LIIIRHKDNIARIKNKTENLVPWG 203


RESULT    7
ID   YF09_HELPY     STANDARD;      PRT;   262 AA.
AC   O26039;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   HYPOTHETICAL PROTEIN HP1509.
GN   HP1509.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE; 97394467.
RA   TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
RA   FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA   NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA   LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
RA   MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA   BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA   COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,
RA   HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA   VENTER J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000649; AAD08551.1; -.
DR   TIGR; HP1509; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM     53     73       POTENTIAL.
FT   TRANSMEM     76     96       POTENTIAL.
FT   TRANSMEM    112    132       POTENTIAL.
FT   TRANSMEM    170    190       POTENTIAL.
FT   TRANSMEM    195    215       POTENTIAL.
SQ   SEQUENCE   262 AA;  28835 MW;  5CE61FE3 CRC32;

  Query Match           16.6%;  Score 256;  DB 1;  Length 262;
  Best Local Similarity  33.1%;  Pred. No. 1.13e-16;
  Matches    54;  Conservative   48;  Mismatches  50;  Indels  11;  Gaps   9;

Db      53 INVIFTLLGYLIGGIPFGYALMKIFYGMDITKIGSGGIGATNVLRALQSKGVSNAKQMAL 112
           | ::: :|:||:|:|| |  :  :|: :::   |||  |:|| :| | :|   :|  ||
Qy       2 ITIVLLILAYLLGSIPSGLWIGQVFFQINLREHGSGNTGTTNTFRIL-GK---KA-GMAT 56

Db     113 LVLILDLFKGMFAVFLSKLFGLDYSLQWMVAIASILGHCYSPFLNFNGGKGVSTIMGSVV 172
           :| | |:||| :| :|: :| |:     : :: :::|| :: |  |:|||:|:|  | :
Qy      57 FV-I-DFFKGTLATLLPIIFHLQGVSPLIFGLLAVIGHTFPIFAGFKGGKAVATSAGVIF 114

Db     173 LLIPIESLIGLTVWFFVGKVL-KISSLASILG-VGTAT-VLIF 212
           : ||  |  |:: || : |  : ||:|: : :::   ||:|
Qy     115 GFAPIFCLY-LAIIFFGALYLGSMISLSSVTASIAAVIGVLLF 156


RESULT    8
ID   Y247_MYCGE     STANDARD;      PRT;   239 AA.
AC   P47489;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   HYPOTHETICAL PROTEIN MG247.
GN   MG247.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE; 96026346.
RA   FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
RA   FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,
RA   FRITCHMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,
RA   NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
RA   TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,
RA   PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403(1995).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U39703; AAC71467.1; -.
DR   TIGR; MG247; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM      7     27       POTENTIAL.
FT   TRANSMEM     61     81       POTENTIAL.
FT   TRANSMEM     99    119       POTENTIAL.
FT   TRANSMEM    135    155       POTENTIAL.
FT   TRANSMEM    159    179       POTENTIAL.
FT   TRANSMEM    199    219       POTENTIAL.
SQ   SEQUENCE   239 AA;  27489 MW;  2CBEFCEF CRC32;

  Query Match           12.2%;  Score 188;  DB 1;  Length 239;
  Best Local Similarity  33.3%;  Pred. No. 1.75e-08;
  Matches   62;  Conservative   53;  Mismatches  49;  Indels  22;  Gaps  15;

Db       9 ILVIFSLASGYLLGSIIFADI-FSKIL-KKNVREFGSKNPGATNSMRVFGLKIGFLVAIF 66
           | ::: |  :||||||  : : :: ::   |:|| || | |:||: |::| | |    ::
Qy       2 ITIVL-LILAYLLGSIP-SGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVI 59

Db      67 DAFKGFFAFLLTWILFRFGLQGYLTEKVYQSTYFLSYLSCFAATIGHIFPLYFKFKGGKA 126
           | ||| :| ||  |:|:  ||| ::  :: :   |  |    | ||| ||::  ||||||
Qy      60 DFFKGTLATLLP-IIFH--LQG-VSPLIF-G---L--L----AVIGHTFPIFAGFKGGKA 105

Db     127 IATTGGSLLAISLWWFLICLLIWIMITL-ITKYVSLASLITFFVLAV-IILIPWLDYLYF 184
           :||::| :::::   | : | | :  :| :   :||:|: : :   : ::|:| : :: :
Qy     106 VATSAGVIFGFAPI-FCLYLAIIFFGALYLGSMISLSSVTASIAAVIGVLLFPLFGFI-L 163

Db     185 FNSDPL 190
             | |:|
Qy     164 SNYDSL 169


RESULT    9
ID   Y247_MYCPN     STANDARD;      PRT;   239 AA.
AC   P75428;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   HYPOTHETICAL PROTEIN MG247 HOMOLOG.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE; 97105885.
RA   HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
RA   HERRMANN R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE000047; AAB96134.1; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM      6     26       POTENTIAL.
FT   TRANSMEM     62     82       POTENTIAL.
FT   TRANSMEM     99    119       POTENTIAL.
FT   TRANSMEM    135    155       POTENTIAL.
FT   TRANSMEM    159    179       POTENTIAL.
FT   TRANSMEM    199    219       POTENTIAL.
SQ   SEQUENCE   239 AA;  27439 MW;  58287C2D CRC32;

  Query Match           10.2%;  Score 158;  DB 1;  Length 239;
  Best Local Similarity  31.3%;  Pred. No. 4.06e-05;
  Matches   55;  Conservative   53;  Mismatches  47;  Indels  21;  Gaps  13;

Db      11 IVFSLVIGYLMGSVMFADV-FGKIL-NKDVRKLGSKNPGATNSIRVFGLKIGFLVGLCDA 68
           ||: |:::||:||:  : : :| ::   ;:|  || | |:||::|::| | |    : |
Qy       4 IVL-LILAYLLGSIP-SGLWIGQVFFQINLREHGSGNTGTTNTFRILGKKAGMATFVIDF 61

Db      69 LKGFLAFVFSFLIFSFWLQQYLNVNQYQKVYYLTYLSCFAATIGHIFPLYFKFKGGKAIA 128
           :|| || ::: :||   ||    |:    :: |  |    | ||| ||::  ||||||:|
Qy      62 FKGTLATLLP-IIFH--LQ---GVSPL--IFGL--L----AVIGHTFPIFAGFKGGKAVA 107

Db     129 TTGGSLLAISLWWFVICLVLWLLVTL-ITKYVSLASLVTFFILAI-IILVPWLDYL 182
           |::| :::::   | : | : :: :| :   :||:|: : :   | ::| | : ::
Qy     108 TSAGVIFGFAPI-FCLYLAIIFFGALYLGSMISLSSVTASIAAVIGVLLFPLFGFI 162


RESULT   10
ID   NAH1_MOUSE     STANDARD;      PRT;   820 AA.
AC   Q61165;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+) EXCHANGER 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RA   DEWEY M.J., BOWMAN L.H.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U51112; AAA92976.1; -.
DR   MGD; MGI:102462; SLC9A1.
DR   PFAM; PF00999; Na_H_Exchanger; 1.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     12       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       13     32       A (M1) HYDROPHOBIC.
FT   DOMAIN       33    105       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      106    127       B (M2) HYDROPHOBIC.
FT   DOMAIN      128    130       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    131    150       C (M3) (POTENTIAL).
FT   DOMAIN      151    162       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    163    183       D (M4) (POTENTIAL).
```

```
FT   DOMAIN      184    188       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    189    210       E (M5) (POTENTIAL).
FT   DOMAIN      211    230       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    231    251       F (M5A) (POTENTIAL).
FT   DOMAIN      252    260       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    261    282       G (M5B) (POTENTIAL).
FT   DOMAIN      283    301       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    302    322       H (M6) (POTENTIAL).
FT   DOMAIN      323    336       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    337    357       I (M7) (POTENTIAL).
FT   DOMAIN      358    388       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    389    410       J (M8) (POTENTIAL).
FT   DOMAIN      411    416       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    417    438       K (M9) (POTENTIAL).
FT   DOMAIN      439    452       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      453    473       L, HYDROPHOBIC.
FT   DOMAIN      474    482       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    483    503       M (M10) (POTENTIAL).
FT   DOMAIN      504    820       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    374    374       POTENTIAL.
SQ   SEQUENCE   820 AA;  91467 MW;  CB3788F4 CRC32;

  Query Match             9.4%;  Score 145;  DB 1;  Length 820;
  Best Local Similarity  27.3%;  Pred. No. 9.97e-04;
  Matches    36;  Conservative   40;  Mismatches  45;  Indels  11;  Gaps   9;

Db    262 LVFGESLLNDAVTVVLYHLFEEFASYDSVGISD-IFLGFLSFFVVALGGVFVGVVY-GVI 319
          |:||  ||  ||    : :|  | :  :|: |  :::|| ::| : |: :| | :| | :
Qy     82 LIFG--LL--AVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIFCLYLAIIFFGALYLGSM 137

Db    320 AAFTSRFTSHIRVIEPLFVF-LYSYMAYLSAELFHLSGIMALIASGVVMRPYVE-ANISH 377
          :::|  |: |  :  :::| |:::      || :: |:|| || :::|   : | | :
Qy    138 ISLSS-VTASIAAVIGVLLFPLFGFILSNYDSLF-IAIILAL-ASLIIIRHKDNIARIKN 194

Db    378 KSHTTIKYFLKM 389
          |:   :   |::
Qy    195 KTENLVPWGLNL 206


RESULT 11
ID   NAH1_BOVIN     STANDARD;      PRT;   817 AA.
AC   Q28036;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+) EXCHANGER 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae;
OC   Bovinae; Bos.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RA   ZHU H., ZHANG Q., ZHANG X., LIU W., TRUMBLY R.J., GARLID K.D.,
RA   SUN X.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U49432; AAA91483.1; -.
DR   PFAM; PF00999; Na_H_Exchanger; 1.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     15       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     16     35       M1 (POTENTIAL).
FT   DOMAIN       36    107       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    108    127       M2 (POTENTIAL).
FT   DOMAIN      128    129       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    130    149       M3 (POTENTIAL).
FT   DOMAIN      150    154       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    155    174       M4 (POTENTIAL).
FT   DOMAIN      175    191       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    192    211       M5 (POTENTIAL).
FT   DOMAIN      212    227       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    228    247       M5A (POTENTIAL).
FT   DOMAIN      248    256       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    257    276       M5B (POTENTIAL).
FT   DOMAIN      277    294       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    295    315       M6 (POTENTIAL).
FT   DOMAIN      316    338       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    339    358       M7 (POTENTIAL).
FT   DOMAIN      359    386       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    387    406       M8 (POTENTIAL).
FT   DOMAIN      407    410       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    411    430       M9 (POTENTIAL).
FT   DOMAIN      431    480       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    481    500       M10 (POTENTIAL).
FT   DOMAIN      501    817       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     75     75       POTENTIAL.
FT   CARBOHYD    370    370       POTENTIAL.
SQ   SEQUENCE   817 AA;  91017 MW;  42535859 CRC32;

  Query Match             9.2%;  Score 142;  DB 1;  Length 817;
  Best Local Similarity  27.3%;  Pred. No. 2.05e-03;
  Matches    36;  Conservative   37;  Mismatches  48;  Indels  11;  Gaps   9;

Db    258 LVFGESLLNDAVTVVLYHLFEEFANYDRVGI-VDIILGFLSFFVVSLGGVFVGVVY-GVI 315
          |:||  ||  ||    : :|  |      |:    :|:|| ::| : |: :| | :| | :
Qy     82 LIFG--LL--AVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIFCLYLAIIFFGALYLGSM 137

Db    316 AAFTSRFTSHIRVIEPLFVF-LYSYMAYLSAELFHLSGIMALIASGVVMRPYVE-ANISH 373
          :::|  |: |  :  :::| |:::      || :: |:|| || :::|   : | | :
Qy    138 ISLSS-VTASIAAVIGVLLFPLFGFILSNYDSLF-IAIILAL-ASLIIIRHKDNIARIKN 194

Db    374 KSHTTIKYFLKM 385
          |:   :   |::
Qy    195 KTENLVPWGLNL 206


RESULT 12
ID   NAH1_CRIGR     STANDARD;      PRT;   822 AA.
AC   P48761;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+) EXCHANGER 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93192332.
RA   COUNILLON L., POUYSSEGUR J.;
RT   "Nucleotide sequence of the Chinese hamster Na+/H+ exchanger NHE1.";
RL   Biochim. Biophys. Acta 1172:343-345(1993).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
```

```
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X68970; CAA48771.1; -.
DR   PFAM; PF00999; Na_H_Exchanger; 1.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     12       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       13     32       A (M1) HYDROPHOBIC.
FT   DOMAIN       33    105       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      106    127       B (M2) HYDROPHOBIC.
FT   DOMAIN      128    130       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    131    150       C (M3) (POTENTIAL).
FT   DOMAIN      151    162       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    163    183       D (M4) (POTENTIAL).
FT   DOMAIN      184    188       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    189    210       E (M5) (POTENTIAL).
FT   DOMAIN      211    230       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    231    251       F (M5A) (POTENTIAL).
FT   DOMAIN      252    260       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    261    282       G (M5B) (POTENTIAL).
FT   DOMAIN      283    301       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    302    322       H (M6) (POTENTIAL).
FT   DOMAIN      323    336       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    337    357       I (M7) (POTENTIAL).
FT   DOMAIN      358    388       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    389    410       J (M8) (POTENTIAL).
FT   DOMAIN      411    416       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    417    438       K (M9) (POTENTIAL).
FT   DOMAIN      439    452       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      453    473       L, HYDROPHOBIC.
FT   DOMAIN      474    482       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    483    503       M (M10) (POTENTIAL).
FT   DOMAIN      504    822       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    374    374       POTENTIAL.
SQ   SEQUENCE   822 AA;  92003 MW;  17073216 CRC32;

  Query Match             9.2%;  Score 142;  DB 1;  Length 822;
  Best Local Similarity  26.5%;  Pred. No. 2.05e-03;
  Matches    35;  Conservative   40;  Mismatches  46;  Indels  11;  Gaps   9;

Db    262 LVFGESLLNDAVTVVLYHLFEEFANYDSIGISD-IFLGFLSFFVVALGGVFVGVVY-GVI 319
          |:||  ||  ||    : :|  |    ::: |  :::|| ::| : |: :| | :| | :
Qy     82 LIFG--LL--AVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIFCLYLAIIFFGALYLGSM 137

Db    320 AAFTSRFTSHIRVIEPLFVF-LYSYMAYLSAELFHLSGIMALIASGVVMRPYVE-ANISH 377
           :::|  |: |  :  :::| |::::      || :: |:|| || :::|   : | | :
Qy    138 ISLSS-VTASIAAVIGVLLFPLFGFILSNYDSLF-IAIILAL-ASLIIIRHKDNIARIKN 194

Db    378 KSHTTIKYFLKM 389
          |:   :   |::
Qy    195 KTENLVPWGLNL 206


RESULT   13
ID   NAH1_RABIT     STANDARD;      PRT;   816 AA.
AC   P23791;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-DEC-1999 (Rel. 39, Last annotation update)
DE   SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+) EXCHANGER 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW ZEALAND WHITE; TISSUE=ILEAL VILLUS;
RX   MEDLINE; 91293066.
RA   TSE C.-M., MA A.I., YANG V.W., WATSON A.J.M., LEVINE S.,
RA   MONTROSE M.H., POTTER J., SARDET C., POUYSSEGUR J., DONOWITZ M.;
RT   "Molecular cloning and expression of a cDNA encoding the rabbit ileal
RT   villus cell basolateral membrane Na+/H+ exchanger.";
RL   EMBO J. 10:1957-1967(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 92096447.
RA   HILDEBRANDT F., PIZZONIA J.H., REILLY R.F., REBOUCAS N.A.,
RA   SARDET C., POUYSSEGUR J., SLAYMAN C.W., ARONSON P.S., IGARASHI P.;
RT   "Cloning, sequence, and tissue distribution of a rabbit renal Na+/H+
RT   exchanger transcript.";
RL   Biochim. Biophys. Acta 1129:105-108(1991).
RN   [3]
RP   SEQUENCE OF 472-816 FROM N.A.
RC   STRAIN=NEW ZEALAND WHITE; TISSUE=HEART MUSCLE;
RX   MEDLINE; 91138752.
RA   FLIEGEL L., SARDET C., POUYSSEGUR J., BARR A.;
RT   "Identification of the protein and cDNA of the cardiac Na+/H+
RT   exchanger.";
RL   FEBS Lett. 279:25-29(1991).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: KIDNEY AND INTESTINE.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X59935; CAA42558.1; -.
DR   EMBL; X61504; CAA43721.1; -.
DR   EMBL; X56536; CAA39881.1; -.
DR   PIR; S13926; S13926.
DR   PIR; S16328; S16328.
DR   PFAM; PF00999; Na_H_Exchanger; 1.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     15       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     16     35       M1 (POTENTIAL).
FT   DOMAIN       36    107       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    108    127       M2 (POTENTIAL).
FT   DOMAIN      128    129       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    130    149       M3 (POTENTIAL).
FT   DOMAIN      150    154       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    155    174       M4 (POTENTIAL).
FT   DOMAIN      175    191       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    192    211       M5 (POTENTIAL).
```

```
FT   DOMAIN      212    227       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    228    247       M5A (POTENTIAL).
FT   DOMAIN      248    256       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    257    276       M5B (POTENTIAL).
FT   DOMAIN      277    294       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    295    315       M6 (POTENTIAL).
FT   DOMAIN      316    338       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    339    358       M7 (POTENTIAL).
FT   DOMAIN      359    386       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    387    406       M8 (POTENTIAL).
FT   DOMAIN      407    410       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    411    430       M9 (POTENTIAL).
FT   DOMAIN      431    480       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    481    500       M10 (POTENTIAL).
FT   DOMAIN      501    816       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     75     75       POTENTIAL.
FT   CARBOHYD    370    370       POTENTIAL.
FT   CONFLICT    242    242       V -> A (IN REF. 2).
FT   CONFLICT    569    569       K -> E (IN REF. 2).
SQ   SEQUENCE   816 AA;  90717 MW;  6315C4BF CRC32;

  Query Match             9.1%;  Score 141;  DB 1;  Length 816;
  Best Local Similarity  27.3%;  Pred. No. 2.61e-03;
  Matches    36;  Conservative   37;  Mismatches  48;  Indels  11;  Gaps   9;

Db    258 LVFGESLLNDAVTVVLYHLFEEFANYDHVGI-VDIVLGFLSFFVVALGGVFVGVVY-GVI 315
          |:||  ||  ||    : :|  |     |:    :::|| ::| : |: :| | :| | :
Qy     82 LIFG--LL--AVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIFCLYLAIIFFGALYLGSM 137

Db    316 AAFTSRFTAHIRVIEPLFVF-LYSYMAYLSAELFHLSGIMALIASGVVMRPYVE-ANISH 373
          :::|  || |  :  :::| |:::      || :: |:|| || :::|   : | | :
Qy    138 ISLSS-VTASIAAVIGVLLFPLFGFILSNYDSLF-IAIILAL-ASLIIIRHKDNIARIKN 194

Db    374 KSHTTIKYFLKM 385
          |:   :   |::
Qy    195 KTENLVPWGLNL 206


RESULT   14
ID   NAH1_PIG       STANDARD;      PRT;   818 AA.
AC   P48762;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+) EXCHANGER 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92087905.
RA   REILLY R.F., HILDEBRANDT F., BIEMESDERFER D., SARDET C.,
RA   POUYSSEGUR J., ARONSON P.S., SLAYMAN C.W., IGARASHI P.;
RT   "cDNA cloning and immunolocalization of a Na(+)-H+ exchanger in
RT   LLC-PK1 renal epithelial cells.";
RL   Am. J. Physiol. 261:F1088-F1094(1991).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M89631; AAA31092.1; -.
DR   EMBL; S71135; AAB20633.1; -.
DR   PFAM; PF00999; Na_H_Exchanger; 1.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     11       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       12     31       A (M1) HYDROPHOBIC.
FT   DOMAIN       32    101       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      102    123       B (M2) HYDROPHOBIC.
FT   DOMAIN      124    126       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    127    146       C (M3) (POTENTIAL).
FT   DOMAIN      147    158       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    159    179       D (M4) (POTENTIAL).
FT   DOMAIN      180    184       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    185    206       E (M5) (POTENTIAL).
FT   DOMAIN      207    226       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    227    247       F (M5A) (POTENTIAL).
FT   DOMAIN      248    256       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    257    278       G (M5B) (POTENTIAL).
FT   DOMAIN      279    297       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    298    318       H (M6) (POTENTIAL).
FT   DOMAIN      319    332       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    333    353       I (M7) (POTENTIAL).
FT   DOMAIN      354    384       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    385    406       J (M8) (POTENTIAL).
FT   DOMAIN      407    412       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    413    434       K (M9) (POTENTIAL).
FT   DOMAIN      435    448       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      449    469       L, HYDROPHOBIC.
FT   DOMAIN      470    478       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    479    499       M (M10) (POTENTIAL).
FT   DOMAIN      500    818       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    370    370       POTENTIAL.
FT   CONFLICT    683    683       H -> Y (IN AAB20633).
SQ   SEQUENCE   818 AA;  90987 MW;  771B7051 CRC32;

  Query Match             9.1%;  Score 140;  DB 1;  Length 818;
  Best Local Similarity  26.5%;  Pred. No. 3.31e-03;
  Matches    35;  Conservative   38;  Mismatches  48;  Indels  11;  Gaps   9;

Db    258 LVFGESLLNDAVTVVLYHLFEEFANYDRVGI-VDIVLGFLSFFVVSLGGVFVGVVY-GVI 315
          |:||  ||  ||    : :|  |     |:    :::|| ::| : |: :| | :| | :
Qy     82 LIFG--LL--AVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIFCLYLAIIFFGALYLGSM 137

Db    316 AAFTSRFTSHIRVIEPLFVF-LYSYMAYLSAELFHLSGIMALIASGVVMRPYVE-ANISH 373
          :::|  |: |  :  :::| |:::   :  || :: |:|| || :::|   : | | :
Qy    138 ISLSS-VTASIAAVIGVLLFPLFGFILSNYDSLF-IAIILAL-ASLIIIRHKDNIARIKN 194

Db    374 KSHTTIKYFLKM 385
          |:   :   |::
Qy    195 KTENLVPWGLNL 206


RESULT   15
ID   NAH1_RAT       STANDARD;      PRT;   820 AA.
AC   P26431;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+) EXCHANGER 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=HEART;
RX   MEDLINE; 92250539.
```

```
RA   ORLOWSKI J., KANDASAMY R.A., SHULL G.E.;
RT   "Molecular cloning of putative members of the Na/H exchanger gene
RT   family. cDNA cloning, deduced amino acid sequence, and mRNA tissue
RT   expression of the rat Na/H exchanger NHE-1 and two structurally
RT   related proteins.";
RL   J. Biol. Chem. 267:9331-9339(1992).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: NOT TISSUE SPECIFIC.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M85299; AAA98479.1; -.
DR   PIR; A40204; A40204.
DR   PFAM: PF00999; Na_H_Exchanger; 1.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     12       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       13     32       A (M1) HYDROPHOBIC.
FT   DOMAIN       33    105       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      106    127       B (M2) HYDROPHOBIC.
FT   DOMAIN      128    130       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    131    150       C (M3) (POTENTIAL).
FT   DOMAIN      151    162       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    163    183       D (M4) (POTENTIAL).
FT   DOMAIN      184    188       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    189    210       E (M5) (POTENTIAL).
FT   DOMAIN      211    230       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    231    251       F (M5A) (POTENTIAL).
FT   DOMAIN      252    260       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    261    282       G (M5B) (POTENTIAL).
FT   DOMAIN      283    301       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    302    322       H (M6) (POTENTIAL).
FT   DOMAIN      323    336       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    337    357       I (M7) (POTENTIAL).
FT   DOMAIN      358    388       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    389    410       J (M8) (POTENTIAL).
FT   DOMAIN      411    416       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    417    438       K (M9) (POTENTIAL).
FT   DOMAIN      439    452       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      453    473       L, HYDROPHOBIC.
FT   DOMAIN      474    482       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    483    503       M (M10) (POTENTIAL).
FT   DOMAIN      504    820       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    374    374       POTENTIAL.
SQ   SEQUENCE   820 AA;  91647 MW;  5286E511 CRC32;

  Query Match            9.0%;  Score 139;  DB 1;  Length 820;
  Best Local Similarity  27.3%;  Pred. No. 4.20e-03;
  Matches    36;  Conservative   39;  Mismatches  46;  Indels  11;  Gaps   9;

Db    262 LVFGESLLNDAVTVVLYHLFEEFASYEYVGISD-IFLGFLSFFVVSLGGVFVGVVY-GVI 319
          |:||  ||  ||    : :|  | :   |: |  :::|| ::| : |: :| | :| | : .
Qy     82 LIFG--LL--AVIGHTFPIFAGFKGGKAVATSAGVIFGFAPIFCLYLAIIFFGALYLGSM 137

Db    320 AAFTSRFTSHIRVIEPLFVF-LYSYMAYLSAELFHLSGIMALIASGVVMRPYVE-ANISH 377
          :::|  |: |  :  :::| |::::      || :: |:|| || :::|   : | | :

Qy    138 ISLSS-VTASIAAVIGVLLFPLFGFILSNYDSLF-IAIILAL-ASLIIIRHKDNIARIKN 194

Db    378 KSHTTIKYFLKM 389
          |:   :   |::
Qy    195 KTENLVPWGLNL 206


Search completed: Wed Mar  8 10:23:31 2000
Job time : 14 secs.
```